



Trinocular Geometry Revisited

Matthew Trager, Jean Ponce, Martial Hebert

► To cite this version:

Matthew Trager, Jean Ponce, Martial Hebert. Trinocular Geometry Revisited. International Journal of Computer Vision, 2016. hal-01152348v3

HAL Id: hal-01152348

<https://hal.science/hal-01152348v3>

Submitted on 15 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Trinocular Geometry Revisited

Matthew Trager · Jean Ponce · Martial Hebert

Received: date / Accepted: date

Abstract When do the visual rays associated with triplets of point correspondences converge, that is, intersect in a common point? Classical models of trinocular geometry based on the fundamental matrices and trifocal tensor associated with the corresponding cameras only provide partial answers to this fundamental question, in large part because of underlying, but seldom explicit, general configuration assumptions. This paper uses elementary tools from projective line geometry to provide necessary and sufficient geometric and analytical conditions for convergence in terms of transversals to triplets of visual rays, without any such assumptions. In turn, this yields a novel and simple minimal parameterization of trinocular geometry for cameras with non-collinear or collinear pinholes, which can be used to construct a practical and efficient method for trinocular geometry parameter estimation. We present numerical experiments using synthetic and real data.

Keywords Trinocular geometry · Trilinearities · Minimal parameterizations of camera geometry · Camera parameter estimation.

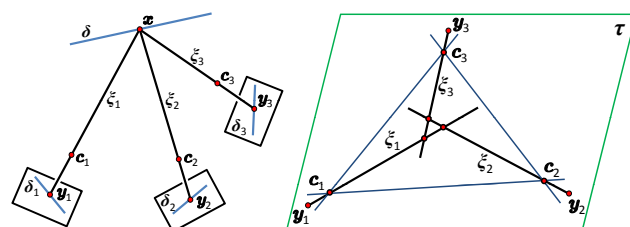


Fig. 1 Left: Visual rays associated with three (correct) correspondences. Right: Degenerate epipolar constraints associated with three coplanar, but non-intersecting rays lying in the trifocal plane τ (as in the rest of this presentation, the image planes are omitted for clarity in this part of the figure). See text for details. Reprinted from Ponce and Hebert (2014).

1 Introduction

The images of points recorded by multiple cameras may only match when the corresponding visual rays *converge*—that is, intersect in a common point (Figure 1, left). For two views, this condition is captured by the bilinear *epipolar constraint* and the corresponding *fundamental matrix* (Longuet-Higgins, 1981; Luong and Faugeras, 1996). Three images can be characterized by both the pairwise epipolar constraints associated with any two of the pictures, and a set of *trilinearities* associated with all three views and parameterized by the associated *trifocal tensor* (Hartley, 1997; Shashua, 1995; Spetsakis and Aloimonos, 1990; Weng et al., 1992). For cameras with non-collinear pinholes, at least, the rays associated with three image points that satisfy the corresponding epipolar constraints almost always converge: The *only* exception is when the points have been matched incorrectly, *and* all lie in the trifocal plane spanned by the three pinholes (Figure 1, right). Interestingly, Hartley and Zisserman state that the fundamental matrices associated with three cameras with non-collinear pinholes determine the corresponding trifocal tensor (Hartley and Zisserman, 2000, Result 14.5), while Faugeras and Mour-

Matthew Trager
Inria, Willow team, ENS/CNRS/Inria UMR 8548
E-mail: matthew.trager@inria.fr.

Jean Ponce
Ecole Normale Supérieure / PSL Research University
Willow team, ENS/CNRS/Inria UMR 8548
E-mail: jean.ponce@ens.fr.

Martial Hebert
Robotics Institute, Carnegie-Mellon University
E-mail: hebert@ri.cmu.edu.

rain (1995) and Ponce et al. (2005), for example, note that the rays associated with three points only satisfying certain (and different) subsets of the trilinearities *alone* must intersect. These claims contradict each other, since rays that satisfy epipolar constraints do not *always* converge, but they are true under some *general configuration* assumptions, rarely made explicit. It is thus worth clarifying these assumptions, and understanding exactly how much the trifocal constraints add to the epipolar ones for point correspondences. This is the problem addressed in this paper, using elementary projective line geometry. In particular, our analysis shows that exploiting both the epipolar constraints and one or two of the trinocular ones, depending on whether the camera pinholes are collinear or not, always guarantees the convergence of the corresponding visual rays. Our analysis also provides, in both cases, a novel and simple minimal parameterization of trinocular geometry.

1.1 Related Work

Geometric constraints involving multiple perspective views of the same point (Figure 1, left) have been studied in computer vision since the seminal work of Longuet-Higgins, who proposed in 1981 the essential matrix as a bilinear model of epipolar constraints between two calibrated cameras (Longuet-Higgins, 1981). Its uncalibrated counterpart, the fundamental matrix, was introduced by Luong and Faugeras (1996). The trilinear constraints associated with three views of a straight line were discovered by Spetsakis and Aloimonos (1990) and by Weng et al. (1992). The uncalibrated case was tackled by Shashua (1995) and by Hartley (1997), who coined the term trifocal tensor. The quadrifocal tensor was introduced by Triggs (1995), and Faugeras and Mourrain (1995) gave a simple characterization of all multilinear constraints associated with multiple perspective images of a point. The usual formulation of the trilinear constraints associated with three images of the same point are asymmetric, one of the images playing a privileged role. A simple and symmetric formulation based on line geometry was introduced in Ponce et al. (2005). A few minimal parameterizations of trinocular geometry are also available (Canterakis, 2000; Papadopoulos and Faugeras, 1998; Ressler, 2002; Torr and Zisserman, 1997).

From a historical point of view, it is worth noting, like Forsyth and Ponce (2003), that epipolar constraints were already known by photogrameters long before they were (re)discovered by Longuet-Higgins (1981), as witnessed by the 1966 Manual of Photogrammetry (Thompson et al., 1966). Although this book does not mention trilinear constraints, it discusses higher-order trinocular *scale-restraint condition equations*: To be more precise, the three fundamental *condition equations* of analytical photogrammetry are identified in (Thompson et al., 1966) as the *collinearity equations* (the

image point y_i , the scene point x , and the pinhole c_i of Figure 1 (left) must belong to the same line ξ_i for $i = 1, 2, 3$), the *coplanarity equations* (the epipolar constraints that express the fact that the lines ξ_1, ξ_2 and ξ_3 intersect pairwise, or equivalently, are pairwise coplanar), and the *scale-restraint equations* (the points where the rays associated with three images of the same point are the closest in space must have the same depth).

The trifocal constraints are explicitly discussed in the latest (2004) edition of the Manual of Photogrammetry (McGlone, 2004), with the aim to express the fact that the three rays ξ_1, ξ_2 , and ξ_3 converge. However, the direct derivation of trifocal constraints for point correspondences typically amounts to writing that all 4×4 minors of some $k \times 4$ matrix are zero, thus guaranteeing that the three lines intersect (Faugeras and Mourrain, 1995; Ponce et al., 2005): The 9×4 matrix used by Faugeras and Mourrain (1995) is obtained by stacking linear combinations of rows of the projection matrices associated with three cameras, whereas the 12×9 matrix of Ponce et al. (2005) is obtained by stacking the *join matrices* that will be defined in Section 2.2.1, and are associated with three visual rays.

These determinants are then rewritten as linear combinations of *reduced minors* that are bilinear or trilinear functions of the image point coordinates (Appendix I gives the corresponding derivation for the case of Ponce et al. (2005) for completeness). The whole difficulty lies in selecting an appropriate subset of reduced minors that will always guarantee that the rays intersect. We have already observed that the bilinear epipolar constraints, alone, are not sufficient. As further discussed in Appendix II, we are not aware of any *fixed* set of four trilinearities that, alone, guarantee convergence in all cases. This suggests seeking instead appropriate combinations of bilinear and trilinear constraints, which is the approach taken in this presentation.

1.2 Problem Statement and Proposed Approach

As noted earlier, the goal of this paper is to understand exactly how much the trifocal constraints add to the epipolar ones for point correspondences. Since both types of constraints model incidence relationships among the light rays joining the cameras' pinholes to observed points, we address this problem using the tools of projective geometry (Veblen and Young, 1910) in general, and line geometry (Pottmann and Wallner, 2001) in particular. As noted earlier, the trifocal tensor was originally invented to characterize the fact that three image lines δ_1, δ_2 , and δ_3 are the projections of the same scene line δ (Shashua, 1995; Spetsakis and Aloimonos, 1990; Weng et al., 1992) (Figure 1, left). The trilinearities associated with three image points y_1, y_2 , and y_3 were then obtained by constructing lines δ_1, δ_2 , and δ_3 passing through these points, and whose preimage is a line δ passing through

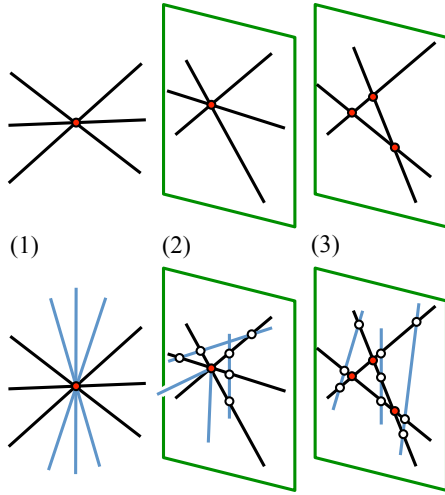


Fig. 2 Top: The possible configurations of three pairwise-coplanar distinct lines, classified according to the way they intersect. The three given lines are shown in black; the planes where two of them intersect are shown in green; and the points where two of the lines intersect are shown in red. Bottom: Transversals to the three lines, shown in blue, and forming (1) a line bundle; (2) a degenerate congruence; and (3) a line field. Reprinted from Ponce and Hebert (2014).

the corresponding scene point x . By construction, this line is a *transversal* to the three rays ξ_1 , ξ_2 , and ξ_3 , that is, it intersects them. It is therefore not surprising that much of the presentation will be dedicated to the characterization of the set of transversals to a triplet of lines.

In particular, we have already seen that the fact that three lines intersect pairwise is necessary, but not sufficient for these lines to intersect. We will show in the rest of this presentation that a necessary and sufficient condition for three lines to converge is in fact that they be pairwise coplanar *and* admit a well defined family of transversals. We will also give a simple geometric and analytical characterization of these transversals under various assumptions. When applied to camera systems, it will provide in turn a new and simple minimal parameterization of trinocular geometry.

Contributions The main contributions of this paper can be summarized as follows:

- We give a new *geometric* characterization of triplets of converging lines in terms of transversals to these lines (Proposition 1).
- We provide a novel and simple *analytical* characterization of triplets of converging lines (Lemma 3 and Proposition 2), that does not rely on the assumptions of general configuration implicit in Ponce et al. (2005).
- We show by applying these results to camera geometry that the three epipolar constraints and one of the trifocal ones (two if the pinholes are collinear) are necessary and sufficient for the corresponding optical rays to converge (Propositions 3 and 4).
- We introduce a new analytical parameterization of epipolar

and trifocal constraints, leading to a *minimal* parameterization of trinocular geometry (Propositions 5 and 6).

- We present an algorithm for estimating trifocal geometry parameters. Our approach minimizes the distance of image points to epipolar and “trinocular” lines, and can be used to reconstruct camera matrices efficiently (not requiring the introduction of auxiliary variables, contrary to bundle adjustment).

The rest of this presentation is organized as follows: Section 2.1 characterizes the manner in which three lines may intersect (or not) each other in geometric terms. This characterization is translated into analytical terms in Section 2.2 using the tools of elementary projective geometry, notably Plücker coordinates and the join matrix (Pottmann and Wallner, 2001; Veblen and Young, 1910). It is then applied to the specific case of epipolar and trifocal constraints in Section 3, where it is used to derive a new minimal parameterization of trinocular geometry (see Canterakis (2000); Papadopoulos and Faugeras (1998); Torr and Zisserman (1997)). An algorithm for recovering the corresponding parameters from at least six point correspondences observed in three images is presented in Section 4, where it is compared to competing methods on both synthetic and real data. We conclude with a brief discussion in Section 5. Formal proofs and technical details are given in three appendices.

A preliminary version of this paper appeared in Ponce and Hebert (2014), without appendices and the full experimental evaluation of Section 4. A few technical improvements to Ponce and Hebert (2014) also include:

- The minimal parameterizations given in Propositions 5 and 6 are based on new sets of constraints, which are generally less susceptible to degeneracies (see Appendix III).
- The “trinocular lines” used in the practical implementation are guaranteed to be well defined and transversal to the epipolar lines applying an appropriate homography of \mathbb{P}^3 (see the discussion in Section 4).

2 Converging Triplets of Lines

2.1 Geometric Point of View

All lines considered from now on are assumed to be different from each other. A *transversal* to some family of lines is a line intersecting every element of this family. We prove in this section the following main result.

Proposition 1 *A necessary and sufficient condition for three lines to converge is that they be pairwise coplanar, and that they admit a transversal not contained in the planes defined by any two of them.*

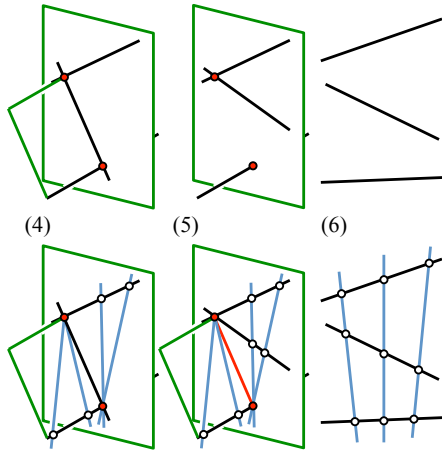


Fig. 3 Top: The possible configurations of three distinct, non-pairwise-coplanar lines, classified according to the way they intersect. Bottom: Transversals to the three lines, forming (4) two pencils of lines having one of the input lines (in black) in common (5) two pencils of lines having one line (in red) in common; and (6) a non-degenerate regulus. See text for details. Reprinted from Ponce and Hebert (2014).

To prove Proposition 1, we need two intermediate results. In projective geometry, two straight lines are either skew to each other or coplanar, in which case they intersect in exactly one point. Our first lemma enumerates the possible incidence relationships among three lines.

Lemma 1 *Three distinct lines can be found in exactly six configurations (Figures 2 and 3, top): (1) the three lines are not all coplanar and intersect in exactly one point; (2) they are coplanar and intersect in exactly one point; (3) they are coplanar and intersect pairwise in three different points; (4) exactly two pairs of them are coplanar (or, equivalently, intersect); (5) exactly two of them are coplanar; or (6) they are pairwise skew.*

The proof is by enumeration. Lemma 1 has an immediate, important corollary—that is, when three lines are pairwise coplanar, either they are not coplanar and intersect in one point (case 1); they are coplanar and intersect in one point (case 2); or they are coplanar, and intersect pairwise in three different points (case 3). In particular, epipolar constraints are satisfied for triplets of (incorrect) correspondences associated with images of points in the trifocal plane containing the pinholes of three non-collinear cameras.

To go further, it is useful to introduce a notion of linear (in)dependence among lines. The geometric definition of independence of lines matches the usual algebraic definition of linear independence, in which, given a coordinate system, a necessary and sufficient for k lines to be linearly dependent is that some nontrivial linear combination of their Plücker coordinate vectors (Section 2.2.1) be the zero vector of \mathbb{R}^6 . Geometrically, the lines linearly dependent on three skew lines form a *regulus* (Veblen and Young, 1910). A regulus is either a *line field*, formed by all lines in a plane; a *line bundle*,

formed by all lines passing through some point; the union of all lines belonging to two flat pencils lying in different planes but sharing one line; or a *non-degenerate regulus* formed by one of the two sets of lines ruling a hyperboloid of one sheet or a hyperbolic paraboloid. Armed with these definitions, we obtain an important corollary of Lemma 1.

Lemma 2 *Three distinct lines always admit an infinity of transversals, that can be found in exactly six configurations (Figures 2 and 3, bottom): (1) the transversals form a bundle of lines; (2) they form a degenerate congruence consisting of a line field and of a bundle of lines; (3) they form a line field; (4) they form two pencils of lines having one of the input lines in common; (5) they form two pencils of lines having a line passing through the intersection of two of the input lines in common; or (6) they form a non-degenerate regulus, with the three input lines in the same ruling, and the transversals in the other one.*

Lemma 2 should not come as a surprise since the transversals to three given lines satisfy three linear constraints and thus form in general a rank-3 family (the degenerate congruence is a rank-4 exception (Veblen and Young, 1910)). Without additional assumptions, not much more can be said in general, since Lemma 2 tells us that any three distinct lines admit an infinity of transversals. When the lines are, in addition, pairwise coplanar, cases 4 to 6 in Lemma 2 are eliminated, and we obtain Proposition 1 as an immediate corollary of this lemma.

2.2 Analytical Point of View

2.2.1 Preliminaries

To translate the geometric results of the previous section into analytical ones, it is necessary to recall a few basic facts about projective geometry in general, and line geometry in particular. Readers familiar with Plücker coordinates, the join operator, etc., may safely proceed to Section 2.2.2. Given some choice of coordinate system for some two-dimensional projective space \mathbb{P}^2 , points and lines in \mathbb{P}^2 can be identified with their homogeneous coordinate vectors in \mathbb{R}^3 . In addition, if \mathbf{x} and \mathbf{y} are two distinct points on a line ξ in \mathbb{P}^2 , we have $\xi = \mathbf{x} \times \mathbf{y}$. A necessary and sufficient condition for a point \mathbf{x} to lie on a line ξ is $\xi \cdot \mathbf{x} = 0$, and two lines intersect in exactly one point or coincide. A necessary and sufficient conditions for three lines to intersect is that they be linearly dependent, or $\text{Det}(\xi_1, \xi_2, \xi_3) = 0$.

In three dimensions, given any choice of coordinate system for a three-dimensional projective space \mathbb{P}^3 , we can identify any line in \mathbb{P}^3 with its Plücker coordinate vector $\xi = (\mathbf{u}; \mathbf{v})$ in \mathbb{R}^6 , where \mathbf{u} and \mathbf{v} are vectors of \mathbb{R}^3 , and we use a semicolon to indicate that the coordinates of \mathbf{u} and \mathbf{v}

have been stacked onto each other to form a vector in \mathbb{R}^6 . In addition, if \mathbf{x} and \mathbf{y} are two distinct points on some line $\xi = (\mathbf{u}; \mathbf{v})$ in \mathbb{P}^3 , we have

$$\mathbf{u} = \begin{bmatrix} x_4 y_1 - x_1 y_4 \\ x_4 y_2 - x_2 y_4 \\ x_4 y_3 - x_3 y_4 \end{bmatrix}, \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{bmatrix}. \quad (1)$$

A Plücker coordinate vector is only defined up to scale, and its \mathbf{u} and \mathbf{v} components are by construction orthogonal to each other—this is sometimes known as the *Klein constraint* $\mathbf{u} \cdot \mathbf{v} = 0$. Let us consider the symmetric bilinear form $\mathbb{R}^6 \times \mathbb{R}^6 \rightarrow \mathbb{R}$ associating with two elements $\lambda = (\mathbf{a}; \mathbf{b})$ and $\mu = (\mathbf{c}; \mathbf{d})$ of \mathbb{R}^6 the scalar $(\lambda|\mu) = \mathbf{a} \cdot \mathbf{d} + \mathbf{b} \cdot \mathbf{c}$. A necessary and sufficient for a nonzero vector ξ in \mathbb{R}^6 to represent a line is that $(\xi|\xi) = 0$, and a necessary and sufficient condition for two lines λ and μ to be coplanar (or, equivalently, to intersect) is that $(\lambda|\mu) = 0$.

Let us denote the basis points of some arbitrary projective coordinate system by \mathbf{x}_0 to \mathbf{x}_4 , with coordinates $\mathbf{x}_0 = (0, 0, 0, 1)^T$, $\mathbf{x}_1 = (1, 0, 0, 0)^T$, $\mathbf{x}_2 = (0, 1, 0, 0)^T$, $\mathbf{x}_3 = (0, 0, 1, 0)^T$, and $\mathbf{x}_4 = (1, 1, 1, 1)^T$. Points \mathbf{x}_0 to \mathbf{x}_3 are called the *fundamental points*. The point \mathbf{x}_4 is the *unit point*. Let us also define four *fundamental planes* \mathbf{p}_j ($j = 0, 1, 2, 3$) whose coordinate vectors are the same as those of the fundamental points. The unique line joining two distinct points is called the *join* of these points and it is denoted by $\mathbf{x} \vee \mathbf{y}$. Likewise, the unique plane defined by a line $\xi = (\mathbf{u}; \mathbf{v})$ and some point \mathbf{x} not lying on this line is called the *join* of ξ and \mathbf{x} , and it is denoted by $\xi \vee \mathbf{x}$. Algebraically, we have $\xi \vee \mathbf{x} = [\xi_\vee] \mathbf{x}$, where $[\xi_\vee]$ is the *join matrix* defined by

$$[\xi_\vee] = \begin{bmatrix} [\mathbf{u}_\times] & \mathbf{v} \\ -\mathbf{v}^T & 0 \end{bmatrix}. \quad (2)$$

A necessary and sufficient condition for a point \mathbf{x} to lie on a line ξ is that $\xi \vee \mathbf{x} = 0$.

2.2.2 Back to Transversals

Let us translate some of the geometric incidence constraints derived in the previous section into algebraic ones. We assume that some projective coordinate system is given, and identify points, planes, and lines with their homogeneous coordinate vectors. Let us consider three distinct lines $\xi_j = (\xi_{1j}, \dots, \xi_{6j})^T$ ($j = 1, 2, 3$), and define

$$D_{ijk} = \begin{vmatrix} \xi_{i1} & \xi_{i2} & \xi_{i3} \\ \xi_{j1} & \xi_{j2} & \xi_{j3} \\ \xi_{k1} & \xi_{k2} & \xi_{k3} \end{vmatrix} \quad (3)$$

to be the 3×3 minor of the 6×3 matrix $[\xi_1, \xi_2, \xi_3]$ corresponding to its rows i, j , and k . A necessary and sufficient condition for this matrix to have rank 2, and thus for the three lines to form a flat pencil (Section 2.1), is that all the minors $T_0 = D_{456}$, $T_1 = D_{234}$, $T_2 = D_{315}$, and $T_3 = D_{126}$ be equal to zero.

Lemma 3 *Given some integer j in $\{0, 1, 2, 3\}$, a necessary and sufficient condition for ξ_1, ξ_2 , and ξ_3 to admit a transversal passing through \mathbf{x}_j is that $T_j = 0$.*

Proof Let us prove the result in the case $j = 0$. The proofs for the other cases are similar. A necessary and sufficient condition for a line $\delta = (\mathbf{u}; \mathbf{v})$ to pass through \mathbf{x}_0 is that $\mathbf{v} = 0$ (this follows from the form of the join matrix). Thus a necessary and sufficient condition for the existence of a line δ passing through \mathbf{x}_0 and intersecting the lines $\xi_j = (\mathbf{u}_j; \mathbf{v}_j)$ is that there exists a vector $\mathbf{u} \neq 0$ such that $(\xi_j|\delta) = \mathbf{v}_j \cdot \mathbf{u} = 0$ for $j = 1, 2, 3$, or, equivalently, that the determinant $T_0 = D_{456} = |\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3|$ be zero.

Combining Proposition 2 and Lemma 3 now yields the following important result.

Proposition 2 *A necessary and sufficient condition for three lines ξ_1, ξ_2 , and ξ_3 to converge is that $(\xi_i|\xi_j) = 0$ for all $i \neq j$ in $\{1, 2, 3\}$, and that $T_j = 0$ for all j in $\{0, 1, 2, 3\}$.*

Proof The condition is clearly necessary. To show that is sufficient, note that since the three lines are pairwise coplanar, they either intersect in exactly one point (cases 1 and 2 of Lemma 2), or are all coplanar, intersecting pairwise in three distinct points, with all their transversals in the same plane (case 3). But the latter case is ruled out by Lemma 3 and the condition $T_j = 0$ for $j = 0, 1, 2, 3$ since the fundamental points \mathbf{x}_j are by construction not all coplanar, and at least one of them (and thus the corresponding transversal) does not lie in the plane containing the three lines.

3 Converging Triplets of Visual Rays

3.1 Bilinearities or Trilinearities?

Let us now turn our attention from general systems of lines to the visual rays associated with three cameras. As noted earlier, it follows from Lemma 1 that the epipolar constraints alone do not ensure that the corresponding viewing rays intersect (Figure 1, right). On the other hand, the only cases where they do not are 1) if the pinholes are not collinear, when the corresponding rays lie in the trifocal plane 2) if the pinholes are collinear, when the visual rays lie in any plane containing the line joining the pinholes. (Figure 4).

As shown in Appendix II (Proposition A), and contrary to the claim of Ponce et al. (2005, Appendix), the trilinear conditions $T_j = 0$ ($j = 0, 1, 2, 3$) associated with three visual rays do not guarantee, on their own, that the rays intersect: In fact, one can always construct a two-dimensional family of triplets of non-intersecting visual rays passing through three given non-collinear pinholes and satisfying these constraints. More generally, we are not aware of any *fixed* set of four trilinearities that can always be chosen to ensure the convergence

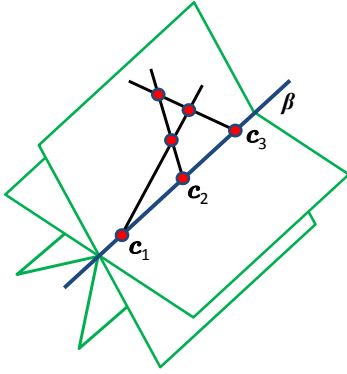


Fig. 4 Degenerate epipolar constraints associated with three images when the three pinholes are collinear and the rays are coplanar but don't intersect in a common point. Reprinted from Ponce and Hebert (2014).

of the corresponding visual rays, which in turns appears to contradict Faugeras and Mourrain (1995, Sec. 4.2.2) (also the discussion in Heyden and Astrom (1997)). This apparent contradiction stems from the fact that both Faugeras and Mourrain (1995) and Ponce et al. (2005) characterize the convergence of visual rays by the vanishing of certain trilinear *reduced minors* of a $k \times 4$ matrix, and have to (implicitly at times) resort to general configuration assumptions to select a representative set of minors. Characterizing the convergence of triplets of lines directly in terms of *both* binocular and trinocular constraints, as in Proposition 2, avoids this difficulty.

3.2 Bilinearities and Trilinearities

By definition, for any choice of projective coordinate system, the four fundamental points x_j ($j = 0, 1, 2, 3$) are not coplanar. When the three pinholes are not collinear, it is thus always possible to choose a projective coordinate system such that one of the fundamental points, say x_0 , does not lie in the trifocal plane, and we obtain the following immediate corollary of Proposition 2.

Proposition 3 *Given three cameras with non-collinear pinholes c_1 , c_2 , and c_3 , and any projective coordinate system such that x_0 does not belong to the trifocal plane, a necessary and sufficient for the three rays $\xi_j = c_j \vee y_j$ ($j = 1, 2, 3$) to converge is that $(\xi_i | \xi_j) = 0$ for all $i \neq j$ in $\{1, 2, 3\}$, and $T_0 = 0$.*

When the three pinholes are collinear (but of course distinct), the three cameras admit a single pencil of epipolar planes, and three rays in epipolar correspondence are in fact always coplanar (Figure 4). The trifocal constraints are necessary in this case to ensure that the three lines intersect in exactly one point. Note that, given three cameras with collinear pinholes, one can always choose a projective coordinate system such that the two fundamental points x_0 and

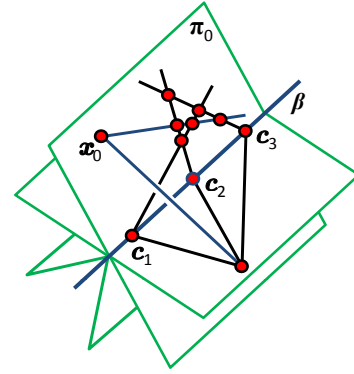


Fig. 5 For collinear pinholes, there exists a single scene plane π_0 in the pencil passing through the baseline β that contains x_0 and for which the condition $T_0 = 0$ is ambiguous. Reprinted from Ponce and Hebert (2014).

x_j (for any j in $\{1, 2, 3\}$) and the baseline joining the three pinholes are not coplanar. The following result characterizes the fact that visual rays intersect in this setting.

Proposition 4 *Given three cameras with collinear pinholes c_1 , c_2 , and c_3 , and any projective coordinate system such that the fundamental points x_0 and x_1 and the baseline β joining the pinholes are not coplanar, a necessary and sufficient condition for the three rays $\xi_j = c_j \vee y_j$ ($j = 1, 2, 3$) to intersect is that $(\xi_i | \xi_j) = 0$ for all $i \neq j$ in $\{1, 2, 3\}$, and $T_0 = T_j = 0$ for some $j \neq 0$.*

Proof The condition is clearly necessary. Because of the epipolar constraints, the three rays must be coplanar, and either intersect in three distinct points with all their transversals in the same plane, intersect in a single point, or coincide with the baseline. Unless the point x_0 lies in the plane π_0 that contains the rays (Figure 5), the first case is ruled out by the condition $T_0 = 0$. If x_0 lies in π_0 , x_j does not (by construction), and the first case is ruled out by $T_j = 0$.

3.3 Epipolar and trinocular constraints

3.3.1 Non-Collinear Pinholes

We assume in this section that the three pinholes are not aligned. In this case, we can always choose a projective coordinate system such that the three fundamental points distinct from x_0 are the three camera centers—that is, $c_j = x_j$ for $j = 1, 2, 3$, and x_0 does not lie in the trifocal plane.

With our choice of coordinate system, and the notation $y_j = (y_{1j}, y_{2j}, y_{3j}, y_{4j})^T$, the three epipolar constraints can be written as

$$\begin{aligned} (x_1 \vee y_1 | x_2 \vee y_2) &= 0 & y_{41}y_{32} &= y_{31}y_{42} \\ (x_1 \vee y_1 | x_3 \vee y_3) &= 0 & \iff y_{41}y_{23} &= y_{21}y_{43} \\ (x_2 \vee y_2 | x_3 \vee y_3) &= 0 & y_{42}y_{13} &= y_{12}y_{43} \end{aligned} \quad (4)$$

Given these constraints, we know from Proposition 2 that a necessary and sufficient conditions for the three visual rays to intersect is that $T_0 = 0$ (the other three trilinearities are trivially satisfied with our choice of coordinate system), which is easily rewritten in our case as

$$y_{21}y_{32}y_{13} = y_{31}y_{12}y_{23}. \quad (5)$$

Note that $y_{4j} = 0$ if and only if \mathbf{y}_j lies in \mathbf{p}_0 , which is also the trifocal plane in our case. As expected, it follows immediately from Eqs. (4) and (5) that, unless $y_{41} = y_{42} = y_{43} = 0$, that is, the observed point lies in the trifocal plane, the epipolar constraints imply the trifocal ones. We now need to translate Eqs. (4-5) to the corresponding equations in *image coordinates*. Let us denote by Π_j ($j = 1, 2, 3$) the 4×3 matrix formed by the coordinate vectors of the basis points for the retinal plane of camera number j . The position of an image point with coordinate vector \mathbf{u}_j in that basis is thus $\mathbf{y}_j = \Pi_j \mathbf{u}_j$. Let us denote by π_{ij}^T the i th row of the matrix Π_j , and use superscripts to index coordinates, i.e., for $k = 1, 2, 3$, π_{ij}^k denotes the k th coordinate of π_{ij} .

Proposition 5 *Given three cameras with non-collinear pinholes and hypothetical point correspondences \mathbf{u}_1 , \mathbf{u}_2 , and \mathbf{u}_3 , a necessary and sufficient condition for the three corresponding rays to converge is that*

$$\begin{aligned} \mathbf{u}_1^T F_{12} \mathbf{u}_2 &= 0 & F_{12} &= \pi_{41} \pi_{32}^T - \pi_{31} \pi_{42}^T \\ \mathbf{u}_1^T F_{13} \mathbf{u}_3 &= 0 & \text{where } F_{13} &= \pi_{41} \pi_{23}^T - \pi_{21} \pi_{43}^T, \text{ and} \\ \mathbf{u}_2^T F_{23} \mathbf{u}_3 &= 0 & F_{23} &= \pi_{42} \pi_{13}^T - \pi_{12} \pi_{43}^T \end{aligned} \quad (6)$$

$$(\pi_{21} \cdot \mathbf{u}_1)(\pi_{32} \cdot \mathbf{u}_2)(\pi_{13} \cdot \mathbf{u}_3) = (\pi_{31} \cdot \mathbf{u}_1)(\pi_{12} \cdot \mathbf{u}_2)(\pi_{23} \cdot \mathbf{u}_3). \quad (7)$$

Moreover, for generic image reference frames (see Appendix III Lemma C1), we can assume that $\pi_1 = (\pi_{21}; \pi_{31}; \pi_{41})$, $\pi_2 = (\pi_{12}; \pi_{32}; \pi_{42})$, and $\pi_3 = (\pi_{13}; \pi_{23}; \pi_{43})$, satisfy the 6 homogeneous constraints

$$\begin{aligned} \pi_{21}^3 &= 0, & \pi_{32}^3 &= 0, & \pi_{13}^3 &= 0, \\ \pi_{31}^3 &= \pi_{41}^3, & \pi_{12}^3 &= \pi_{42}^3, & \pi_{23}^3 &= \pi_{43}^3. \end{aligned} \quad (8)$$

and are thus defined by three groups of 7 coefficients, each one uniquely determined up to a separate scale. This is a minimal, 18dof parameterization of trinocular geometry.

Proof Equations (6) and (7) are obtained immediately by substitution in Eqs. (4) and (5). Together, they provide a 24dof parameterization of the trifocal geometry by the three vectors $\pi_j = (\pi_{1j}; \pi_{2j}; \pi_{3j})$ ($j = 1, 2, 3$), each defined up to scale in \mathbb{R}^9 by 8 independent parameters. Locating the camera pinholes at the fundamental points \mathbf{x}_j ($j = 1, 2, 3$) freezes 9 of the 15 degrees of freedom of the projective ambiguity of projective structure from motion. It is possible to exploit the remaining 6 degrees of freedom, and to impose the constraints of Eq. (8) on the vectors π_j .

Indeed, the general form of a projective transform Q mapping the three fundamental points \mathbf{x}_j onto themselves has 7 coefficients defined up to scale. Applying such a transform to the matrices Π_j ($j = 1, 2, 3$) defined in some arbitrary projective coordinate system, and writing that the matrices $Q\Pi_j$ must satisfy the constraints of Eq. (8) yields a system of 6 homogeneous equations in the 7 nonzero entries of Q . Note that we can generate many different sets of homogeneous constraints by choosing different sets of entries of the vectors π_1 , π_2 , and π_3 . It can be shown that there is always some choice for which the system defining Q admits a unique solution defined up to scale, and that this solution is nonsingular, thus defining a valid change of coordinates. Details of the proof can be found in Appendix III (Proposition C). Together, Eqs. (6), (7) and (8) provide us with a minimal, 18dof parameterization of the trinocular geometry by the three vectors π_1 , π_2 and π_3 now each defined up to scale in \mathbb{R}^9 by only 6 independent parameters.

To the best of our knowledge, the minimal parameterization of trinocular geometry proposed by Papadopoulos and Faugeras (1998) is the only other one known so far to be one-to-one and parametric (other minimal ones, e.g., Canterakis (2000); Torr and Zisserman (1997), impose algebraic constraints). Contrary to Papadopoulos and Faugeras (1998), our parameterization does not require the use of a computer algebra system to impose rank constraints (see Papadopoulos and Faugeras (1998) for details). In addition, our parameterization is symmetric, none of the cameras playing a privileged role.

Let us close this section by noting that Eq. (7) has an interesting geometric interpretation: Any point with coordinate vector \mathbf{u}_1 in the first image that matches points with coordinate vectors \mathbf{u}_2 and \mathbf{u}_3 in the other two, must satisfy (7) and thus belong to the “trinocular line” (our terminology):

$$\tau_1 = [(\pi_{32} \cdot \mathbf{u}_2)(\pi_{13} \cdot \mathbf{u}_3)]\pi_{21} - [(\pi_{12} \cdot \mathbf{u}_2)(\pi_{23} \cdot \mathbf{u}_3)]\pi_{31}. \quad (9)$$

This should not come as a surprise since classical trifocal geometry is defined in terms of line correspondences, and Eq. (7) merely expresses the fact that the image point \mathbf{y}_1 lies on the projection τ_1 of the line τ_0 passing through \mathbf{x}_0 that intersects the rays passing through the other two image points, \mathbf{y}_2 and \mathbf{y}_3 . What is less well known is that the lines τ_1 belong to the pencil generated by the lines π_{21} and π_{31} , which intersect at the point $\mathbf{z}_1 = \pi_{21} \times \pi_{31}$ of the first image, i.e., the projection of \mathbf{x}_0 (and indeed, the line τ_1 is not defined if either \mathbf{u}_2 or \mathbf{u}_3 are images of \mathbf{x}_0). The same reasoning applies to the other two images.

3.3.2 Collinear Pinholes

Let us now assume that the three pinholes are collinear (as noted in McGlone (2004), this case may be important in

practice, in aerial photography for example). Let us position the two pinholes c_1, c_2 in x_1 and x_2 , and the third pinhole, c_3 , in $x_1 + x_2$. We are free to do this since this amounts to choosing c_1 and c_2 as the fundamental points of the baseline joining the three pinholes, and c_3 as its unit point. From Eq. (4):

$$y_{41}y_{32} = y_{31}y_{42}, y_{41}y_{33} = y_{31}y_{43}, y_{42}y_{33} = y_{32}y_{43}, \quad (10)$$

and write $T_0 = 0$ and $T_3 = 0$ respectively as

$$\begin{aligned} y_{31}y_{32}(y_{23} - y_{13}) + y_{33}(y_{31}y_{12} - y_{21}y_{32}) &= 0, \\ y_{41}y_{42}(y_{23} - y_{13}) + y_{43}(y_{41}y_{12} - y_{21}y_{42}) &= 0. \end{aligned} \quad (11)$$

The other two minors T_1 and T_2 are zero with our choice of coordinate system.

We can rewrite as before Eqs. (10) and (11) in terms of the rows of the matrices Π_j ($j = 1, 2, 3$). Given the special role of $y_{23} - y_{13}$ in Eq. (11), it is convenient to introduce the vector $\omega_3 = \pi_{23} - \pi_{13}$, and we obtain the following characterization of the trinocular geometry.

Proposition 6 *Given three cameras with collinear pinholes and hypothetical point correspondences u_1, u_2 , and u_3 , a necessary and sufficient condition for the three corresponding rays to converge is that*

$$\begin{aligned} u_1^T F_{12} u_2 &= 0 & F_{12} &= \pi_{41} \pi_{32}^T - \pi_{31} \pi_{42}^T \\ u_1^T F_{13} u_3 &= 0 & \text{where } F_{13} &= \pi_{41} \pi_{33}^T - \pi_{31} \pi_{43}^T, \\ u_2^T F_{23} u_3 &= 0 & F_{23} &= \pi_{42} \pi_{33}^T - \pi_{32} \pi_{43}^T \end{aligned} \quad (12)$$

$$\begin{aligned} 0 &= (\pi_{31} \cdot u_1)(\pi_{32} \cdot u_2)(\omega_3 \cdot u_3) + \\ &(\pi_{33} \cdot u_3)[(\pi_{31} \cdot u_1)(\pi_{12} \cdot u_2) - (\pi_{21} \cdot u_1)(\pi_{32} \cdot u_2)], \\ 0 &= (\pi_{41} \cdot u_1)(\pi_{42} \cdot u_2)(\omega_3 \cdot u_3) + \\ &(\pi_{43} \cdot u_3)[(\pi_{41} \cdot u_1)(\pi_{12} \cdot u_2) - (\pi_{21} \cdot u_1)(\pi_{42} \cdot u_2)]. \end{aligned} \quad (13)$$

Moreover, for generic image reference frames (see Appendix III Lemma D1) we can assume that $\pi_1 = (\pi_{21}; \pi_{31}; \pi_{41})$, $\pi_2 = (\pi_{12}; \pi_{32}; \pi_{42})$ and $\pi_3 = (\omega_3; \pi_{33}; \pi_{43})$ satisfy the 8 homogeneous constraints

$$\begin{aligned} \pi_{21}^3 &= 0, \pi_{12}^3 = 0, \omega_3^1 = \omega_3^2 = \omega_3^3 = \pi_{33}^3 = \pi_{43}^3 \\ \pi_{31}^3 &= 0, \pi_{42}^3 = 0, \end{aligned} \quad (14)$$

and are thus defined by three groups of, respectively, 6, 6, and 7 independent coefficients, each uniquely determined up to a separate scale, for a total of 16 independent parameters. This is a minimal, 16dof trinocular parameterization.

Proof Equations (12) and (13) are obtained immediately by substitution in Eqs. (10) and (11). Together they provide a 24dof parameterization of the trifocal geometry by the three vectors π_j ($j = 1, 2, 3$), each defined up to scale in \mathbb{R}^9 by 8 independent parameters. Locating the camera pinholes in x_1, x_2 , and $x_1 + x_2$ freezes 7 of the 15 degrees of freedom of the projective ambiguity of projective structure from motion.

Similar to the proof of Proposition 5, the remaining 8 degrees of freedom can be used to impose the constraints of Eq. (14) on the vectors π_j (details can be found in Appendix III, Proposition D). Together, Eqs. (12), (13) and (14) provide us with a minimal, 16dof parameterization of the trinocular geometry by the three vectors π_1, π_2 and π_3 now each defined up to scale in \mathbb{R}^9 by only 5, 5, and 6 independent parameters.

As in the case of non-collinear pinholes, the two trilinear constraints (13) can be seen to express the geometric condition that corresponding points in different images lie on two “trinocular lines”. One can also verify that, in a fixed image, these two sets of trinocular lines belong to the pencils of lines through the projections of x_0 and x_3 .

4 Implementation and Results

Propositions 5 and 6 can be used to estimate the trinocular geometry, specifically the vectors π_j associated with three cameras, from at least six triplets of correspondences between three images. In turn, this information can be used to recover three-dimensional structure using triangulation.

We propose to minimize a “trinocular-epipolar” error, given by the mean squared distance to the epipolar and trinocular lines (9) for all of the image points (cf. below for details). Initial values for the matrices Π_j can be obtained by estimating the projection matrices from six triplets of matching points, using a projective model (Carlsson, 1995; Quan, 1995; Ponce et al., 1994). We have implemented this method, for both non-collinear and collinear pinholes, and compared it with bundle adjustment, and with a baseline that considers only distances to epipolar lines, in order to measure the contribution of trilinear constraints in practice. The next sections contain our experimental results.

We wish to recall that bundle adjustment boils down to the minimization of a “geometric distance error” (Hartley and Zisserman, 2000), which measures the mean squared distance of each given triplet to the nearest image points which actually correspond, according to the current estimate of the camera parameters¹. For example, the contribution of a triplet of image points u_1, u_2, u_3 , assuming estimated parameters Π_1, Π_2, Π_3 can be written as:

$$\begin{aligned} \delta(u_1, u_2, u_3 | \Pi_1, \Pi_2, \Pi_3) &= \\ \min_{(\hat{u}_1, \hat{u}_2, \hat{u}_3)} \{ &d(u_1, \hat{u}_1)^2 + d(u_2, \hat{u}_2)^2 + d(u_3, \hat{u}_3)^2 \} \end{aligned} \quad (15)$$

where $(\hat{u}_1, \hat{u}_2, \hat{u}_3)$ are triplets that correspond for Π_1, Π_2, Π_3 . Despite being very natural, this error has the disadvantage of not being directly computable from image data, since the

¹ Measuring distances requires a euclidean structure in each image. Clearly, this is not an issue in practice.

nearest corresponding triplet must be determined by solving a separate optimization problem. Indeed, bundle adjustment requires introducing a set of auxiliary variables that represent the coordinates of the original unknown 3D points, which contribute to a total of at least $3n + 18$ degrees of freedom. On the other hand, if we consider the “epipolar error” given by

$$\delta_e(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \mid \Pi_1, \Pi_2, \Pi_3) = \sum_{i=1}^3 d(\mathbf{u}_i, E_{i,i+1}(\mathbf{u}_{i+1}))^2 + d(\mathbf{u}_i, E_{i,i+2}(\mathbf{u}_{i+2}))^2, \quad (16)$$

where $E_{ij}(\mathbf{u})$ represents the epipolar line in the image i corresponding to the point \mathbf{u} in image j and addition is done modulo 3, then δ_e can easily be computed from image coordinates and camera parameters (for example using the expressions in Propositions 5 and 6), so the overall minimization problem only has 18 degrees of freedom (16 for collinear pinholes). As we will see, the “trinocular-epipolar” error defined in the next section can also be computed very efficiently, and we will compare the effectiveness of these three error functions using both synthetic and real data. As it is the most meaningful error function, the geometric distance (15) will generally be used to evaluate the quality of the solutions recovered using the different approaches.

Finally, we should point out that enforcing the constraints for the minimal parameterizations (i.e., reducing the degrees of freedom by imposing that the parameters Π satisfy (8) and (14)) does not affect the accuracy of the reconstructions, but often improves performance. In general, the magnitude of this effect seems to be highly dependent on the dataset and on the initial estimate of the configuration, so in our implementation we focused on comparing various error functions and the quality of the reconstructions, rather than analyzing the impact on efficiency of the minimal parameterizations.

4.1 Non-collinear pinholes

In the case of three cameras with non-collinear pinholes, we define the trinocular-epipolar error as the average squared distance between each image point, the three epipolar lines, and one trinocular line (9), according to Proposition 5. More precisely, the contribution of a triplet $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ is given by

$$\delta_t(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \mid \Pi_1, \Pi_2, \Pi_3) = \sum_{i=1}^3 d(\mathbf{u}_i, E_{i,i+1}(\mathbf{u}_{i+1}))^2 + d(\mathbf{u}_i, E_{i,i+2}(\mathbf{u}_{i+2}))^2 + d(\mathbf{u}_i, T_0^{i+1,i+2}(\mathbf{u}_{i+1}, \mathbf{u}_{i+2}))^2, \quad (17)$$

where addition is always modulo 3, and $T_0^{i+1,i+2}(\mathbf{u}, \mathbf{v})$ is the trinocular line in the i -th image, as defined in (9). Note that

this error function can be easily computed from the expressions in Proposition 5. As previously observed, $T_0^{i+1,i+2}(\mathbf{u}, \mathbf{v})$ is actually a well-defined line if and only if \mathbf{x}_0 does not project onto \mathbf{u} or \mathbf{v} in images j, k . We can guarantee this condition by applying an appropriate homography of \mathbb{P}^3 that brings \mathbf{x}_0 to project (using initial camera parameters) “very far” from the given image data. In practice, we choose \mathbf{x}_0 to project at infinity in the three images (which is always possible).

We first evaluate this approach on synthetic data: we consider 50 random configurations of triplets of camera matrices (simulating realistic extrinsic and intrinsic parameters) and clouds of 20 points (see Figure 6), and successively add various amounts of Gaussian noise to the projected points. The averaged results are shown in Figure 7. We note that the trinocular-epipolar error gives slightly better results than the epipolar one, but bundle adjustment is the most accurate (as one would expect since it actually minimizes the geometric distance error (15)). In other words, including the trilinear constraint brings a slight improvement but does not seem to be essential: indeed, we know that for non-collinear pinholes, and three-dimensional points lying in general position, enforcing the epipolar constraints is sufficient for guaranteeing correspondence. However, if we position the 3D points gradually closer to the trifocal plane, we observe that the epipolar error gives increasingly worse results, ultimately failing in improving the initial reconstruction (Figure 8). In fact, when the epipolar lines are close to being coincident, measuring only epipolar distances can cause severe numerical instabilities. On the other hand, using the trinocular-epipolar error and bundle adjustment, the quality of the reconstruction seems to be independent of the distance to the trifocal plane.²

We also evaluated the trinocular-epipolar approach on real data, using three images from the “house dataset” (courtesy of B. Boufama and R. Mohr). Figure 9 shows 38 correspondences between the images from this dataset, and the corresponding epipolar and trinocular lines, after camera parameters were recovered using the trinocular-epipolar approach. Table 1 shows the average distances between the data points and these lines. The mean distance to epipolar lines is on the order of 1 pixel, and comparable to that obtained by classical techniques for estimating the fundamental matrix from pairs of images on the same data (Forsyth and Ponce, 2012, Ch. 8). Finally, Table 2 compares the results of the different approaches, in terms of the geometric distance error. In this case, we see that the reconstructions obtained using the trinocular-epipolar and the epipolar approach are both very accurate.

² We should point out that we could not place all the points exactly on the trifocal plane, since the minimal reconstruction method we used for initialization requires points to be in general position (Carlsson, 1995).

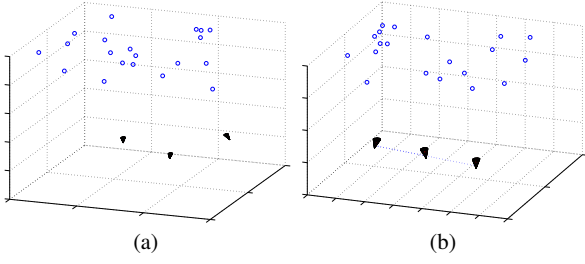


Fig. 6 Examples of synthetic data: non-collinear pinholes (left) and collinear pinholes (right). The camera parameters defined random rotations and translations for camera motions, and produced feasible image sizes (around 500 pixels per dimension).

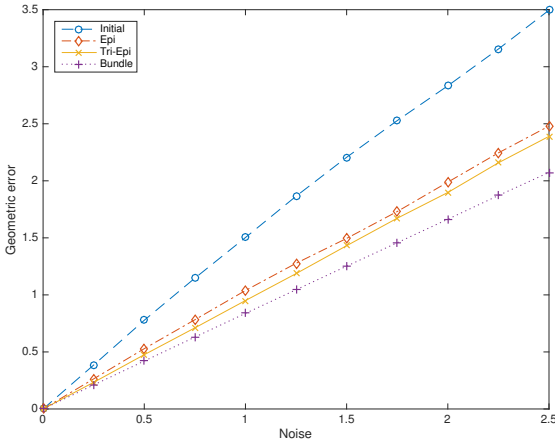


Fig. 7 Non-collinear pinholes: quantitative results using synthetic datasets (20 points) with different amounts of gaussian noise, in terms of the geometric distance error. The different lines represent the error of the initial approximation (Initial), and the errors of the solution refined using the trinocular-epipolar (Tri-Epi), epipolar (Epi) and bundle adjustment (Bundle) approach. The results are averages of 50 configurations of points and cameras.

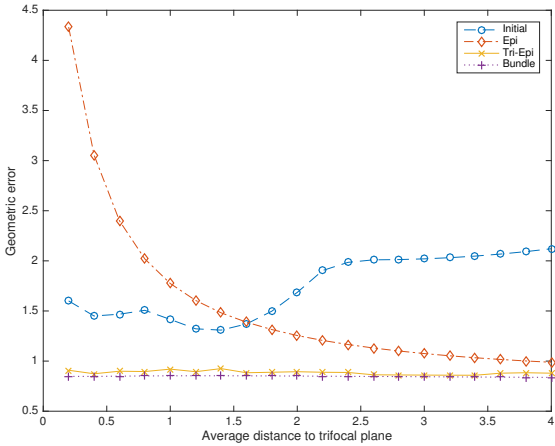


Fig. 8 Near-degenerate datasets: quantitative results using synthetic datasets (20 points) with points lying near the trifocal plane, in terms of the geometric distance error. This experiment shows the same camera configuration and point sets increasingly close to the trifocal plane (constructed as points on the trifocal plane perturbed by gaussian noise).

<i>Epi12</i>	<i>Epi13</i>	<i>Tr1</i>	<i>Epi23</i>	<i>Epi21</i>	<i>Tr2</i>	<i>Epi31</i>	<i>Epi32</i>	<i>Tr3</i>
0.81	0.81	1.03	0.94	0.78	0.79	0.75	0.86	1.28

Table 1 Average distances (in pixels) to epipolar and trinocular lines for the house dataset (Figure 9). Here, “*Epi_{ij}*” refers to the distance between points in image *i* and the corresponding epipolar lines associated with image *j*, and “*Tr_j*” refers to the distance between points in image *j* and the corresponding trinocular line associated with the other two images.

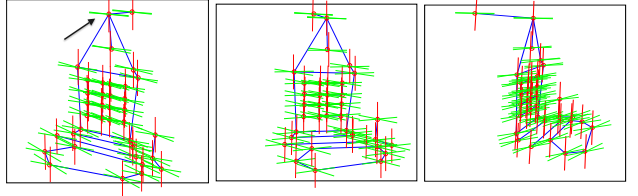


Fig. 9 Estimated epipolar and trinocular lines. Note that the two families of epipolar lines associated with an image typically contain (near) degenerate pairs (such as the pair shown in the first image): these can be disambiguated using trilinearities.

<i>Tri-Epi</i>	<i>Epi</i>	<i>Bundle adj</i>
0.73	0.73	0.72

Table 2 Quantitative comparison between different approaches for the house data-set (Figure 9), in terms of the geometric distance error (measured in pixels).

4.2 Collinear pinholes

For three cameras with collinear pinholes, the “trinocular-epipolar” error is defined as the average squared distance of each image point to the three epipolar lines and to two “trinocular lines”, following Proposition 6:

$$\begin{aligned} \delta_t(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \mid \Pi_1, \Pi_2, \Pi_3) = & \\ & \sum_{i=1}^3 d(\mathbf{u}_i, E_{i,i+1}(\mathbf{u}_{i+1}))^2 + d(\mathbf{u}_i, E_{i,i+2}(\mathbf{u}_{i+2}))^2 \\ & + d(\mathbf{u}_i, T_A^{i+1,i+2}(\mathbf{u}_{i+1}, \mathbf{u}_{i+2}))^2 + d(\mathbf{u}_i, T_B^{i+1,i+2}(\mathbf{u}_{i+1}, \mathbf{u}_{i+2}))^2, \end{aligned} \quad (18)$$

where $T_A^{i+1,i+2}(\mathbf{u}, \mathbf{v})$ and $T_B^{i+1,i+2}(\mathbf{u}, \mathbf{v})$ are the trinocular lines in the *i*-th image, defined by the trilinear constraints in (13). Once again, the trinocular lines are well defined if and only if \mathbf{x}_0 and \mathbf{x}_3 do not project onto the given image points: in order to satisfy this condition, we select two points which project very far from the image data, and apply a homography of \mathbb{P}^3 which maps \mathbf{x}_0 and \mathbf{x}_3 to these points.

Figure 7 shows the results of our experiments using synthetic data (20 points with various amounts of Gaussian noise added to the projections). We see that in this case the epipolar error fails in recovering the correct parameters, while the trinocular-epipolar error is able to achieve essentially optimal solutions. Indeed, we know that in the collinear case the epipolar constraints are never sufficient to guarantee correspondence.

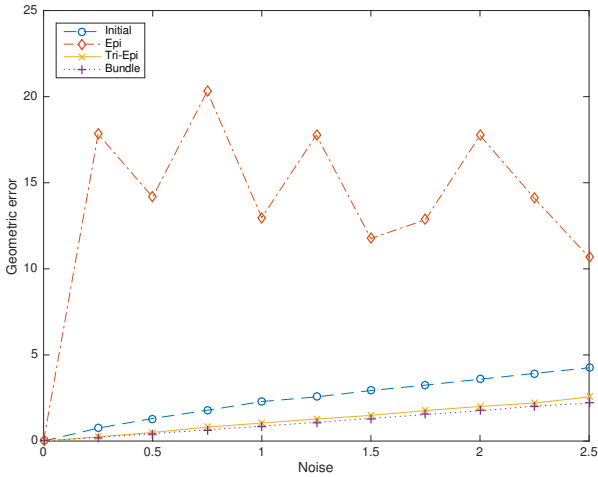


Fig. 10 Collinear pinholes: Quantitative results using synthetic datasets (20 points) with different amounts of gaussian noise, in terms of the geometric distance error. The results are averages of 50 configurations of points and cameras.

Finally, we evaluate our approach on real data, using three pictures taken from collinear viewpoints (allowing rotations about the camera’s axis) and matching points using SIFT descriptors (Figure 11). Table 4 shows the average distances between the data points and the epipolar and trinocular lines for the dataset with collinear-pinholes and Figure 12 displays the estimated epipolar and trinocular lines. Finally, Table 3 shows our quantitative results, which confirm that for collinear pinholes the trilinear components of the error are actually necessary.

Tri-Epi	Epi	Bundle adj
0.68	5.33	0.67

Table 3 Quantitative comparison between different approaches for the office data-set (Figure 11), in terms of the geometric distance error (measured in pixels).

5 Discussion

Our analysis has given us a better understanding of trinocular geometry. In particular, we were able to give necessary and sufficient conditions for point correspondence, based on bilinear and trilinear constraints. To express these conditions analytically we exploited a model for trifocal configurations that led to a family of minimal parameterizations of trinocular geometry, for both non-collinear and collinear pinholes. Although we did not evaluate the impact of such parameterizations on reconstruction algorithms in practice, we believe that these could offer a simple and practical approach to eliminate unnecessary degrees of freedom when optimizing over the set of trinocular structures. We also presented an interesting error function for trifocal parameter estimation, which is based on epipolar and “trinocular” lines, and can be

used to reconstruct camera matrices efficiently (not requiring the introduction of auxiliary variables, contrary to bundle adjustment). Finally, our experiments clearly confirm that enforcing epipolar constraints can correctly recover camera parameters only for non-degenerate configurations, while for collinear pinholes, or for images of 3D points lying close to the trifocal plane, trilinear conditions are necessary.

One may also wonder whether the fact that four lines intersect in exactly one point can also be characterized geometrically or analytically. Indeed, there exists a quadri-focal tensor expressing the corresponding four-view constraints (Triggs, 1995), and it has been shown to be redundant with the epipolar and trifocal constraints. In retrospect, it is geometrically obvious that a necessary and sufficient condition for four lines to intersect in exactly one point is that any two triplets of lines among them also does: this follows immediately from the fact that these triplets have two lines in common, so the point where these two lines intersect is also the point where all four lines intersect. In other words, there is no need to write any equation to realize that considering four lines together instead of a set of triplets does not add anything to the geometric picture in this case. On the other hand, the natural algebraic constraints expressing that four lines are linearly dependent yield a set of quadrilinear constraints. However, one can show that the elements of a rank-3 family of lines do not necessarily intersect in a single point (in general they form a regulus). Thus (this kind of) quadrilinearities, *on their own*, are neither necessary nor sufficient, to characterize the fact that the corresponding visual rays intersect. We should mention that the general problem of providing necessary and sufficient conditions for characterizing point correspondences for an arbitrary number of views has recently been addressed in Trager et al. (2015) using tools from algebraic geometry. Spelling out similar results in terms of analytical line geometry would be very useful, and could be an interesting direction for future work.

Acknowledgments. This work was supported in part by the ERC grant VideoWorld, the Institut Universitaire de France, the Inria - CMU associate team GAYA, and ONR MURI N000141010934.

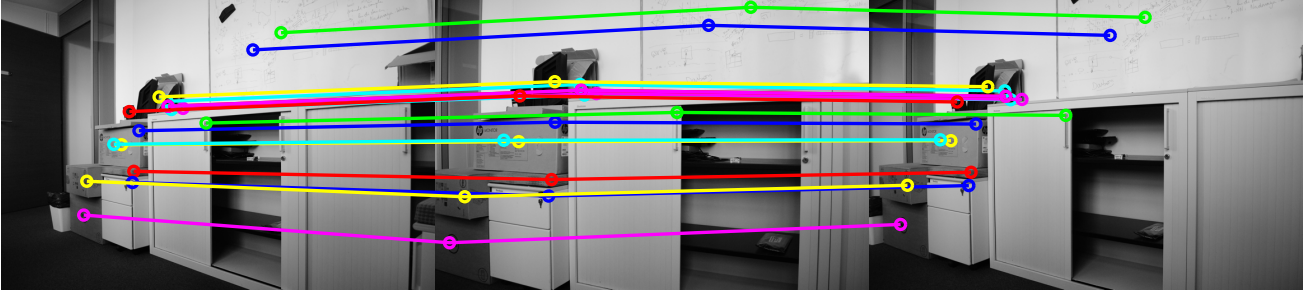


Fig. 11 Three pictures taken from collinear viewpoints, and point correspondences obtained by matching SIFT descriptors.

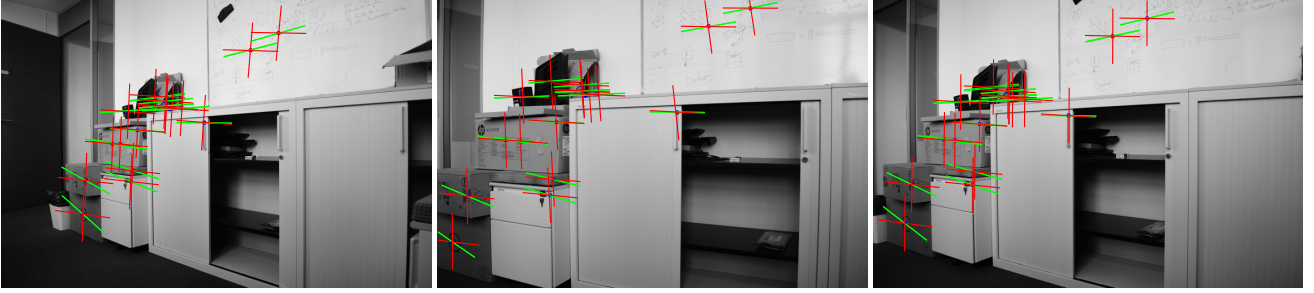


Fig. 12 Estimated epipolar and trinocular lines. Note that the pairs of epipolar lines essentially coincide, while trinocular lines are always transversal.

$Epi12$	$Epi13$	$Tr1A$	$Tr1B$	$Epi23$	$Epi21$	$Tr2A$	$Tr2B$	$Epi31$	$Epi32$	$Tr3A$	$Tr3B$
0.60	0.19	0.49	2.95	0.74	0.76	0.77	0.74	0.21	0.66	2.80	0.84

Table 4 Average distances (in pixels) to epipolar and trinocular lines for the collinear pinhole dataset (Figure 11). Here, “ Epi_{ij} ” refers to the distance between points in image i and the corresponding epipolar lines associated with image j , and “ Tr_{jA} ” and “ Tr_{jB} ” refer to the distances between points in image j and the corresponding two trinocular lines associated with the other two images.

Appendix I: Deriving Trilinearities from the Join Matrix

In this section, we recall for completeness the construction of trilinearities characterizing the convergence of three rays proposed in Ponce et al. (2005). Contrary to the characterization of Proposition 3 in the main body of this paper, this construction is based on an analysis of the join matrices associated with these rays, and it relies on implicit general configuration assumptions that we spell out along the way.

The join matrix $[\xi_\vee]$ associated with a line ξ has rank 2. This is geometrically obvious since the planes passing through a line form a one-dimensional pencil. This also follows immediately from the analytical definition of $[\xi_\vee]$ using the Klein constraint. Let us denote by ξ_i^T ($i = 1, 2, 3$) the first three rows of $[\xi_\vee]$ and by ξ_0^T its fourth row. The rows ξ_i^T ($i = 0, 1, 2, 3$) can be interpreted as the coordinate vectors of four planes in the pencil with axis ξ , such that plane ξ_i contains the fundamental point x_i . This follows from the fact that $\xi_i = [\xi_\vee]^T x_i$. Indeed, since $[\xi_\vee]$ is skew-symmetric, we also have $\xi_i = -[\xi_\vee] x_i = -(\xi \vee x_i)$, i.e., ξ_i contains ξ and x_i . It follows that a necessary and sufficient condition for two rows ξ_i and ξ_j of the join matrix to be linearly independent is that ξ and the line joining x_i and x_j do not intersect.

Now, a necessary and sufficient condition for three lines λ , μ and ν to intersect in a point x is that the homogeneous linear system of 12 equations in 4 unknowns obtained by stacking the corresponding join matrices and defined by

$$\begin{bmatrix} [\lambda_\vee] \\ [\mu_\vee]x \\ [\nu_\vee] \end{bmatrix} = 0 \quad (19)$$

admit a non-trivial solution, or equivalently, that all the 4×4 minors of the corresponding 12×4 matrix be zero.

We will assume in the rest of this section that the lines of interest are in general position relative to the fundamental points, that is, none of them intersects the line joining two fundamental points. It follows that any two rows of the corresponding rank-2 join matrices are linearly independent. Thus, for example any row of $[\lambda_\vee]$ can be written as unique linear combination of, say, $\bar{\lambda}_1^T$ and $\bar{\lambda}_2^T$, and none of the coefficients λ_i is equal to zero.

The 4×4 minors of the 12×4 matrix of Eq. (19) come in two shapes: (1) minors with two rows associated with one of the three lines, and two associated with another one; and (2) minors with two rows associated with one line, and one each from the other two lines. Minors of type (1) only involve two lines and, as shown below, boil down to bilinear constraints. As noted in Ponce et al. (2005) and shown below for completeness, the minors of type (2) are linear combinations of trilinearities.

Let us consider for example the minor associated with the 4×4 matrix formed by any two rows of $[\lambda_\vee]$ and any two rows of $[\mu_\vee]$. Under our assumptions, the first two rows of

this matrix completely determine λ , and the other two completely determine μ , thus a necessary and sufficient condition for the minor to be zero is that these two lines intersect or equivalently $(\lambda|\mu) = 0$. This is easily verified analytically. For example, the minor of the matrix formed by the first two rows of $[\lambda_\vee]$ and the first two rows of $[\mu_\vee]$ is equal to $-\lambda_3\mu_3(\lambda|\mu)$. According to our general position assumption, it follows that a necessary and sufficient condition for this minor to be zero is that λ and μ intersect.

Let us now turn to minors involving two rows from one join matrix, say $[\lambda_\vee]$, and one row each from the other two, and define

$$A_{ijkl} = \frac{1}{\lambda_3} \text{Det} \begin{bmatrix} \bar{\lambda}_i^T \\ \bar{\lambda}_j^T \\ \bar{\mu}_k^T \\ \bar{\nu}_l^T \end{bmatrix}. \quad (20)$$

A simple calculation shows that

$$\begin{aligned} A_{1211} &= -D_{234}, \\ A_{1222} &= -D_{315}, \\ A_{1212} + A_{1221} &= D_{134} - D_{235}, \\ A_{1212} - A_{1221} &= \lambda_3(\mu|\nu) - \mu_3(\nu|\lambda) - \nu_3(\lambda|\mu). \end{aligned} \quad (21)$$

These four “reduced minors” are therefore trilinear in the Plücker coordinates of the three lines. They are in general linearly independent, but they verify some algebraic constraints.

Recall from Section 2.2.1 that $\bar{\delta}_i$ is the plane passing through the line δ and the fundamental point x_i . The minor A_{12kl} thus characterizes the fact that the line where $\bar{\mu}_k$ and $\bar{\nu}_l$ intersect is a transversal to the three lines λ , μ , and ν (Figure 13).

Under our general position assumption, and because of the multilinearity of determinants, any of the reduced minors can be written as a linear combination of $D_{234} = T_1$, $D_{315} = T_2$, A_{1212} , and A_{1221} . This is for example the case of the reduced minors $T_0 = D_{456}$, and $T_3 = D_{126}$, and we can write

$$\begin{aligned} T_0 &= a_0 T_1 + b_0 T_2 + c_0 A_{1212} + d_0 A_{1221}, \\ T_3 &= a_3 T_1 + b_3 T_2 + c_3 A_{1212} + d_3 A_{1221}, \end{aligned} \quad (22)$$

and it follows that all trilinearities can be written as linear combinations of the constraints $T_j = 0$ ($j = 0, 1, 2, 3$) when $c_0 d_3 - c_3 d_0 \neq 0$. As shown in the next section, and contrary to the claim of Ponce et al. (2005), this may not be the case.

Appendix II: Are Trilinearities Alone Sufficient for Convergence?

We show in this section that the constraints $T_j = 0$ ($j = 1, 2, 3$), alone, do not guarantee that the corresponding rays converge (Proposition A). On the other hand, there exist in general four trilinearities among those defined in the previous section that do guarantee convergence (Proposition B).

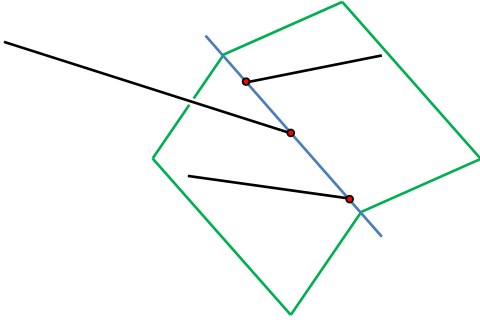


Fig. 13 When a line that intersects two planes that contain two other lines, the intersection (blue line) of these two planes is a transversal to the three lines.

In turn, this implies that all trilinearities cannot always be expressed as linear combinations of the constraints $T_j = 0$ alone.

Trilinearities that don't Guarantee Convergence

Proposition A *Given three pinholes in general position, there exists a two-dimensional family of triples of non-intersecting open neighborhoods of viewing rays going through these pinholes and satisfying the four trilinearities $T_j = 0$ for $j = 0, 1, 2, 3$.*

Proof Let us assume some fixed projective coordinate system. A quadric surface that contains the fundamental points x_i ($i = 0, 1, 2, 3$) satisfies the equation

$$ax_1x_2 + bx_1x_3 + cx_1x_4 + dx_2x_3 + ex_2x_4 + fx_3x_4 = 0. \quad (23)$$

It follows that, given the position of three pinholes c_1, c_2 , and c_3 , there exists in general a non-empty, two-parameter family of quadrics passing through these three points and the four fundamental points, and satisfying the corresponding three instances of Eq. (23) in the six unknowns a to f . This family of quadrics corresponds to a non-empty two-dimensional subspace Q of the projective space \mathbb{P}^5 equipped with homogeneous coordinates (a, b, c, d, e, f) .

Among these quadrics, the (nondegenerate) ruled ones verify the constraint that the associated symmetric matrix

$$A = \begin{bmatrix} 0 & a & b & c \\ a & 0 & d & e \\ b & d & 0 & f \\ c & e & f & 0 \end{bmatrix} \quad (24)$$

has two positive and two negative eigenvalues. Since the matrix is symmetric its eigenvalues are real. Furthermore, they are the roots of the characteristic polynomial

$$\begin{aligned} \lambda^4 - (a^2 + b^2 + c^2 + d^2 + e^2 + f^2)\lambda^2 \\ - 2(abd + ace + bcf + def)\lambda \\ + (a^2f^2 + b^2e^2 + c^2d^2 - 2acdf - 2bcde). \end{aligned} \quad (25)$$

Let a_i denote the coefficient of the monomial of degree i of this polynomial. Since $a_4 = 1$, the product of the roots is $a_0 = a^2f^2 + b^2e^2 + c^2d^2 - 2acdf - 2bcde$, and their sum is $-a_3 = 0$. In particular, a sufficient condition for the quadric to be ruled is that $a_0 > 0$ since, in this case, none of the roots can be zero, and, since their sum is zero, at least one of them is negative and another one positive, forcing two of the roots to be positive and two negative, since otherwise their product would be negative.

Under a general configuration assumption, to be made explicit in a minute, there always exists a nonempty open subset O of Q such that $a_0 > 0$ for any point in O : Indeed, taking $c = 0$ defines a non-empty projective subspace Q' of Q , such that $a_0 = (af - be)^2$ for any point in Q' . In general, Q' is one-dimensional (a straight line), and it either intersects the quadric hypersurface H' defined by $af - be = 0$ in zero, one, or two points (we say that we are in a general configuration), or it is entirely contained in that hypersurface. Assuming we are in a general configuration, the complement of $Q' \cap H'$ in Q' is open, and contains at least one point q where $a_0 > 0$. The value of a_0 for any point in a small enough open neighborhood O of q in Q is also strictly positive. Note that Q' may be of dimension greater than one if some of the equations defining it are redundant, but this only relaxes the general configuration assumption further.

Let us give an example. Assume that the three camera centers are $c_1 = (1, 0, 1, 1)^T$, $c_2 = (0, 1, 1, 1)^T$, and $c_3 = (1, 1, 0, 1)^T$. It is easy to show that the general form of Q in this case is

$$ax_2(x_1 + x_3 - x_4) + bx_3(x_1 + x_2 - x_4) + c[x_4(x_1 - x_2 - x_3) + 2x_2x_3] = 0, \quad (26)$$

which is indeed a two-parameter family of nondegenerate ruled quadrics with $a_0 = 4c(a + c)(b + c)(a + b + c) > 0$ when a, b , and c are, for example, all strictly positive. It is also easy to verify that none of the lines joining two of the optical centers, or two of the fundamental points lie on Q .

The rulings of the quadric $x^T A x$ passing through a given point x have a direction v that satisfies

$$\begin{aligned} x'_1v_1 + x'_2v_2 + x'_3v_3 &= 0, \\ av_1v_2 + bv_1v_3 + dv_2v_3 &= 0, \end{aligned} \quad (27)$$

where $x' = Ax$. Writing $v_1 = -(x'_2v_2 + x'_3v_3)/x'_1$ and substituting in the second equation yields

$$ax'_2v_2^2 + (ax'_3 + bx'_2 - dx'_1)v_2v_3 + bx'_3v_3^2 = 0. \quad (28)$$

This equation can be solved for v_2/v_3 if $|ax'_2| > |bx'_3|$, and for v_3/v_2 otherwise. Note that in our example, we have $d = a + b + 2c$, $e = -a - c$, and $f = -b - c$.

Under our general configuration assumption, there exists a two-dimensional family of nondegenerate ruled quadrics satisfying $a_0 > 0$, and for each one of them, there exists one

line δ_j passing through each pinhole c_j ($j = 1, 2, 3$) in one of the two rulings of the quadric. The lines δ_1, δ_2 , and δ_3 do not intersect each other (unless two of them coincide, which is not the case in our example), but are all intersected by each one of the lines of the second ruling that pass through the four fundamental points (again, no two of these lines coincide in our example), and thus satisfy the corresponding trilinearities (Lemma 3). (Note that a second set of three lines passing through the pinholes and four transversals passing through the fundamental points can be obtained by inverting the roles of the two rulings.)

In other words, trilinearities alone do not guarantee that the associated correspondences are correct.

Figure 14 shows an example, where the three camera centers are $c_1 = (1, 0, 1, 1)^T$, $c_2 = (0, 1, 1, 1)^T$, and $c_3 = (1, 1, 0, 1)^T$. It is easy to show that the general form of a quadric surface passing through these pinholes and the fundamental points is

$$ax_2(x_1 + x_3 - x_4) + bx_3(x_1 + x_2 - x_4) + c[x_4(x_1 - x_2 - x_3) + 2x_2x_3] = 0, \quad (29)$$

which is indeed a two-parameter family of nondegenerate ruled quadrics when a, b , and c are, for example, all strictly positive. It is also easy to verify that none of the lines joining two of the optical centers, or two of the fundamental points lie on Q .

Trilinearities that do Guarantee Convergence

Proposition B *A sufficient condition for three distinct lines λ, μ, ν in general position to converge is that the trilinearities corresponding to the reduced minors $A_{jkjj}, A_{jkjk}, A_{jkkj}$ and A_{jkkk} vanish for some $0 \leq j < k \leq 3$.*

We suppose as before in the rest of this section that our three lines are distinct, and that the line joining the fundamental points x_j and x_k ($0 \leq j < k \leq 3$) does not intersect μ or ν . This implies that any two rows of any of the join matrices are linearly independent, and in particular that they completely determine the corresponding line. This implies as well that the planes $\bar{\mu}_j$ and $\bar{\mu}_k$ (resp. $\bar{\tau}_j$ and $\bar{\tau}_k$) are distinct for $k \neq j$.

Proof Let us first note that the condition $A_{jkjj} = A_{jkkk} = 0$ characterizes the fact that the line λ intersects the planes $\bar{\mu}_j = \mu \vee x_j$ and $\bar{\tau}_j = \nu \vee x_j$. The condition $A_{jkjk} = A_{jkkj} = 0$, on the other hand, characterizes the fact that λ intersects the planes $\bar{\mu}_j = \mu \vee x_j$ and $\bar{\tau}_k = \nu \vee x_k$.

Let us suppose that the lines μ and ν are skew to each other. This implies that the planes $\bar{\mu}_j$ and $\bar{\tau}_k$ are distinct (even if $k = j$) since the lines μ and ν respectively belong to these two planes. It follows that the six lines $\mu, \nu, \delta_j, \delta_k, \delta_{jk}$ and δ_{kj} are the edges of a proper tetrahedron. In particular,

δ_j and δ_k (resp. δ_{jk} and δ_{kj}) are skew to each other. This is the case depicted in Figure 15.

Suppose now that the line λ intersects δ_j, δ_k , and δ_{jk} . Since it intersects δ_j and δ_{jk} , this line must lie in their common plane $\bar{\mu}_j$, or pass through their intersection z_j on the line ν . Likewise, since λ intersects δ_k and δ_{jk} , this line must lie in their common plane $\bar{\tau}_k$ or pass through their intersection z_k on μ .

These four possibilities boil down to three distinct cases: the line λ may either (1) coincide with δ_{jk} (this corresponds to λ lying in both $\bar{\mu}_j$ and $\bar{\tau}_k$ or, equivalently, passing through both z_j and z_k), (2) belong to the flat pencil of lines passing through z_j in the plane $\bar{\tau}_k$, or (3) belong to the flat pencil of lines passing through z_k in the plane $\bar{\mu}_j$.

If we further assume that λ intersects δ_{kj} , case (1) is ruled out by the fact that δ_{jk} and δ_{kj} are skew to each other. Case (2) is ruled out by the fact that the only line in the flat pencil of lines passing through z_j in $\bar{\tau}_k$ that intersect δ_{kj} is ν itself, which is inconsistent with our hypothesis that the three lines λ, μ , and ν are distinct. Case (3) is also ruled out by the fact that the only line in the flat pencil of lines passing through z_k in $\bar{\mu}_j$ that intersect δ_{kj} is μ itself.

Therefore our hypotheses are inconsistent and the two lines μ and ν must be coplanar.

Thus, let us assume from now on that these two lines intersect in some point z (uniquely defined since they are distinct), and examine first the case where $\bar{\mu}_j$ and $\bar{\tau}_k$ coincide (and thus do not define a unique line δ_{jk}). In this case, we have $\delta_j = \bar{\mu}_j \wedge \bar{\tau}_j = \bar{\tau}_k \wedge \bar{\tau}_j = \nu$, $\delta_k = \bar{\mu}_k \wedge \bar{\tau}_k = \bar{\mu}_k \wedge \bar{\mu}_j = \mu$. In particular, a line λ intersecting $\delta_j = \nu$ and $\delta_k = \mu$ must either intersect their common plane $\bar{\mu}_j = \bar{\tau}_k$ transversally, and pass through z (in which case the three lines converge), or lie in that plane.

In the latter case, the three lines λ, μ , and ν are coplanar. Since μ and ν are distinct lines, the planes $\bar{\mu}_k$ and $\bar{\tau}_j$ cannot coincide too, and it follows that the line δ_{kj} is uniquely defined and intersects the plane $\bar{\mu}_j = \bar{\tau}_k$ transversally in z (if it lied in that plane, the two lines μ and ν would coincide with δ_{kj} and each other). Since λ intersects δ_{kj} it must thus pass through z , ensuring the convergence of the three lines λ, μ and ν .

A similar line of reasoning shows that the three lines also converge when $\bar{\mu}_k$ and $\bar{\tau}_j$ coincide, $\bar{\mu}_j$ and $\bar{\tau}_j$ coincide, or $\bar{\mu}_k$ and $\bar{\tau}_k$ coincide. We are thus left with one final case, where the four planes $\bar{\mu}_j, \bar{\mu}_k, \bar{\tau}_j$, and $\bar{\tau}_k$ are pairwise distinct, and the two lines μ and ν intersect in the point z . In this case the four lines $\delta_j, \delta_k, \delta_{jk}$, and δ_{kj} are uniquely defined, and they intersect in z . They do not all lie in the same plane: for example δ_j and δ_{jk} lie in $\bar{\mu}_j$, and δ_j and δ_{kj} lie in $\bar{\tau}_j$. Thus for λ to intersect these four lines, it must also pass through z , and the three lines λ, μ , and ν converge.

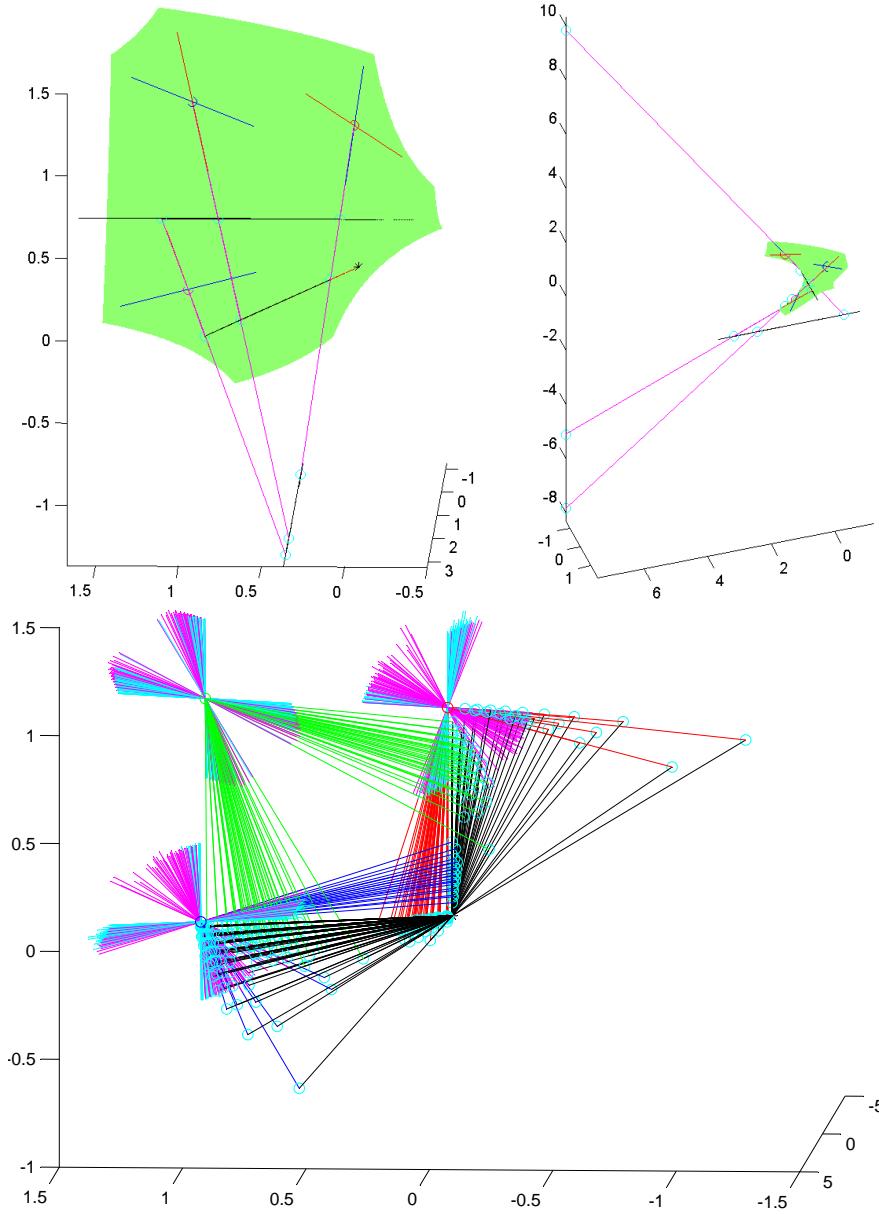


Fig. 14 Ambiguity locus of trilinearities. Top: Three camera rays and their transversals passing through the fundamental points. Bottom: A two-dimensional family of non-intersecting triplets of rays going through the camera centers and satisfying the trilinearities.

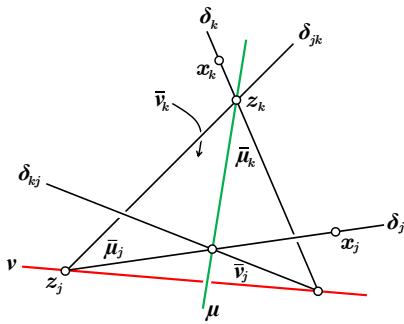


Fig. 15 Transversals corresponding to different minors. See text for details.

Appendix III: Minimal Parameterizations

Non-Collinear Pinholes

For general (non-collinear) pinholes, it is well known and easy to see that the trinocular geometry has 18 degrees of freedom: 11×3 independent elements in the projection matrices, minus 15 degrees from the freedom of choice of projective reference. Considering equations (6) and (7), we notice that none of the constraints involve elements π_{jj} , for $j = 1, 2, 3$, i.e. they don't involve elements on the j -th line of Π_j ; this is due to the fact that “moving” the basis points of the retinal planes along the lines passing through the associated camera

pinholes (effectively changing retinal planes) does not affect the algebraic description of the cameras. Thus, we see that just by positioning the pinholes in the fundamental points $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ we obtain a 24 degree of freedom parameterization; we will use the remaining six degrees of freedom to reduce the parameters to the minimal number.

We will say that three non-collinear cameras are “in general configuration” if no two fundamental points in different retinal planes correspond. This condition obviously does not depend on the projective coordinate system of \mathbb{P}^3 , but one can always apply a general change of coordinates in the three retinal planes in order to guarantee this form of generality.

In the following, we use the same notation introduced in the main body of the paper.

Lemma C1 *Three non-collinear cameras, with pinholes in $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$, are in general configuration if and only if for all permutations (i, j, h) of $(1, 2, 3)$*

$$\det \begin{bmatrix} \pi_{hi}^k & \pi_{4i}^k \\ \pi_{hj}^l & \pi_{4j}^l \end{bmatrix} \neq 0 \quad (30)$$

holds for all $k, l \in \{1, 2, 3\}$.

Proof One easily verifies that the condition $\det \begin{bmatrix} \pi_{hi}^k & \pi_{4i}^k \\ \pi_{hj}^l & \pi_{4j}^l \end{bmatrix} = 0$ expresses the fact that the lines $\mathbf{x}_i \times \Pi_i^k$ and $\mathbf{x}_j \times \Pi_j^l$ intersect.

Proposition C *Given three cameras with non-collinear pinholes in general configuration, it is possible to find a projective coordinate system of \mathbb{P}^3 setting the pinholes in the fundamental points $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$, and such that, up to permutation of the plane coordinates, the basis points of the retinal planes are such that*

$$\begin{aligned} \pi_{21}^3 &= 0, & \pi_{32}^3 &= 0, & \pi_{13}^3 &= 0, \\ \pi_{31}^3 &= \pi_{41}^3, & \pi_{12}^3 &= \pi_{42}^3, & \pi_{23}^3 &= \pi_{43}^3. \end{aligned} \quad (31)$$

In other words, we can assume that the Π_j matrices, for $j = 1, 2, 3$ are of the following form:

$$\Pi_1 = \begin{bmatrix} X & X & X \\ * & * & 0 \\ * & * & 1 \\ * & * & 1 \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} * & * & 1 \\ X & X & X \\ * & * & 0 \\ * & * & 1 \end{bmatrix}, \quad \Pi_3 = \begin{bmatrix} * & * & 0 \\ * & * & 1 \\ X & X & X \\ * & * & 1 \end{bmatrix} \quad (32)$$

where the “*” indicate unconstrained elements, that yield an 18 dof parametrization.

Proof Once we set the pinholes in the fundamental points \mathbf{x}_j ($j = 1, 2, 3$), the general form of a projective transformation Q that maps the points $\mathbf{x}_1, \mathbf{x}_2$, and \mathbf{x}_3 onto themselves is

$$Q = \begin{bmatrix} \alpha & 0 & 0 & \kappa \\ 0 & \beta & 0 & \lambda \\ 0 & 0 & \gamma & \mu \\ 0 & 0 & 0 & \nu \end{bmatrix}, \quad (33)$$

where $\alpha, \beta, \gamma, \nu$ are nonzero scalars. Let us now show that we can indeed find a matrix Q that will transform the matrices Π_j so as to satisfy the constraints of Eq. (31). In the *original* coordinate system, we have

$$Q\Pi_j = \begin{bmatrix} \alpha\pi_{1j}^T + \kappa\pi_{4j}^T \\ \beta\pi_{2j}^T + \lambda\pi_{4j}^T \\ \gamma\pi_{3j}^T + \mu\pi_{4j}^T \\ \nu\pi_{4j}^T \end{bmatrix}, \quad (34)$$

which allows us to rewrite the constraints of Eq. (31) on the *transformed* matrices Π_j as a system of 6 homogeneous equations in the 7 nonzero entries of Q that can be separated into three groups of two:

$$\begin{aligned} \begin{bmatrix} \pi_{12}^3 & \pi_{42}^3 \\ \pi_{13}^3 & \pi_{43}^3 \end{bmatrix} \begin{bmatrix} \alpha \\ \kappa \end{bmatrix} &= \nu \begin{bmatrix} \pi_{42}^3 \\ 0 \end{bmatrix}, \\ \begin{bmatrix} \pi_{23}^3 & \pi_{43}^3 \\ \pi_{21}^3 & \pi_{41}^3 \end{bmatrix} \begin{bmatrix} \beta \\ \lambda \end{bmatrix} &= \nu \begin{bmatrix} \pi_{43}^3 \\ 0 \end{bmatrix}, \\ \begin{bmatrix} \pi_{31}^3 & \pi_{41}^3 \\ \pi_{32}^3 & \pi_{42}^3 \end{bmatrix} \begin{bmatrix} \gamma \\ \mu \end{bmatrix} &= \nu \begin{bmatrix} \pi_{41}^3 \\ 0 \end{bmatrix}. \end{aligned} \quad (35)$$

The 2×2 determinants involved in these equations are nonzero, because of the general configuration assumption. Thus there will be a unique solution defined up to scale. The solution yields nonzero values for $\alpha, \beta, \gamma, \nu$, determining a valid homography defined by the matrix Q , if the elements π_{41}^3, π_{42}^3 and π_{43}^3 are non-zero. This can always be assumed to be true, up to permuting columns of the Π_j matrices, since these must have at least one non-zero element on the fourth row.

Collinear Pinholes

Assuming collinear pinholes, the degrees of freedom for the trinocular geometry reduce to 16: indeed, the first two camera matrices have 22 independent parameters, while the third one only 9 since it’s center is constrained on a line (two conditions); this means $31 - 15 = 16$ independent parameters by removing projective ambiguity. Fixing the camera centers in $\mathbf{x}_1, \mathbf{x}_2$ and $\mathbf{x}_1 + \mathbf{x}_2$ reduces the number of parameters to 21, and we use the remaining degrees of freedom to impose constraints on the entries of the matrices Π_j ($j = 1, 2, 3$).

We say that three cameras with collinear pinholes are in general position if no two basis vectors of the retinal planes give rise to a correspondence *and* the line containing the pinholes does not intersect the third retina plane on the line $x + y + z = 0$. Once again this condition does not depend on the choice of coordinates in \mathbb{P}^3 , but can be assumed true up to a generic change of coordinates in the three retina planes.

Lemma D1 *Three collinear cameras, with pinholes in $\mathbf{x}_1, \mathbf{x}_2$, $\mathbf{x}_1 + \mathbf{x}_2$, are in general configuration if and only if for all permutations (i, j, h) of $(1, 2, 3)$*

$$\det \begin{bmatrix} \pi_{hi}^k & \pi_{4i}^k \\ \pi_{hj}^l & \pi_{4j}^l \end{bmatrix} \neq 0 \quad (36)$$

holds for all $k, l \in \{1, 2, 3\}$ and the matrix

$$\begin{bmatrix} \pi_{33}^1 - \pi_{33}^2 \pi_{43}^1 - \pi_{43}^2 \\ \pi_{33}^2 - \pi_{33}^3 \pi_{43}^2 - \pi_{43}^3 \\ \pi_{33}^3 - \pi_{33}^1 \pi_{43}^3 - \pi_{43}^1 \end{bmatrix} \quad (37)$$

has rank 2.

Proof As for the non-collinear case, the condition

$$\det \begin{bmatrix} \pi_{hi}^k & \pi_{4i}^k \\ \pi_{hj}^l & \pi_{4j}^l \end{bmatrix} = 0 \quad (38)$$

expresses the fact that the lines $\mathbf{x}_i \times \Pi_i^k$ and $\mathbf{x}_j \times \Pi_j^l$ intersect.

Note that the matrix (37) cannot have rank 0, or the third retina plane would contain $\mathbf{c}_3 = \mathbf{x}_1 + \mathbf{x}_2$; if it had rank 1, it would mean that the unique solution (x, y, z) to

$$\begin{bmatrix} \pi_{33}^1 & \pi_{33}^2 & \pi_{33}^3 \\ \pi_{43}^1 & \pi_{43}^2 & \pi_{43}^3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0 \quad (39)$$

satisfies $x + y + z = 0$: indeed, any linear relation between the first two rows of (37) can be used to find a solution of (39) with $x + y + z = 0$. These coordinates represent the point in the the retina plane such that $\mathbf{x}_3 = \mathbf{x}_4 = 0$, i.e., the intersection with the line generated by the camera pinholes.

Proposition D *Given three cameras with collinear pinholes in general configuration, it is possible to find a projective coordinate system of \mathbb{P}^3 setting the pinholes in the fundamental points $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_1 + \mathbf{x}_2$, and such that the basis points of the retina planes are such that*

$$\begin{aligned} \pi_{21}^3 &= 0, \quad \pi_{12}^3 = 0, \quad \omega_3^1 = \omega_3^2 = \omega_3^3 = \pi_{33}^3 = \pi_{43}^3 \\ \pi_{31}^3 &= 0, \quad \pi_{42}^3 = 0, \end{aligned} \quad (40)$$

In other words, we can assume that the Π_j matrices, for $j = 1, 2, 3$ are of the following form:

$$\begin{aligned} \Pi_1 &= \begin{bmatrix} X & X & X \\ * & * & 0 \\ * & * & 0 \\ * & * & 1 \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} * & * & 0 \\ X & X & X \\ * & * & 1 \\ * & * & 0 \end{bmatrix}, \\ \Pi_3 &= \begin{bmatrix} X & X & X \\ X+1 & X+1 & X+1 \\ * & * & 1 \\ * & * & 1 \end{bmatrix} \end{aligned} \quad (41)$$

where the “*” indicate unconstrained elements, that yield a 16 dof parametrization.

Proof Once we set the pinholes in the fundamental points $\mathbf{x}_1, \mathbf{x}_2$ and $\mathbf{x}_1 + \mathbf{x}_2$, the general form of a projective transformation Q that maps these points onto themselves is

$$Q = \begin{bmatrix} \alpha & 0 & \beta & \gamma \\ 0 & \alpha & \delta & \varepsilon \\ 0 & 0 & \lambda & \mu \\ 0 & 0 & \nu & \tau \end{bmatrix}, \quad (42)$$

with 9 parameters defined up to scale, or 8 independent parameters. Let us show that, as before, we can indeed find a matrix Q that will transform the matrices Π_j so as to satisfy the constraints of Eq. (40). In the original coordinate system, we have

$$Q\Pi_j = \begin{bmatrix} \alpha\pi_{1j}^T + \beta\pi_{3j}^T + \gamma\pi_{4j}^T \\ \alpha\pi_{2j}^T + \delta\pi_{3j}^T + \varepsilon\pi_{4j}^T \\ \lambda\pi_{3j}^T + \mu\pi_{4j}^T \\ \nu\pi_{3j}^T + \tau\pi_{4j}^T \end{bmatrix}, \quad (43)$$

which allows us to rewrite the constraints of Eq. (40) on the transformed matrices Π_j as a system of 8 homogeneous equations in the 9 distinct nonzero entries of Q . These equations can be separated into two groups:

$$\begin{aligned} \begin{bmatrix} \pi_{33}^3 - \pi_{33}^2 \pi_{43}^3 - \pi_{43}^2 \\ \pi_{33}^2 - \pi_{33}^1 \pi_{43}^2 - \pi_{43}^1 \\ \pi_{33}^1 - \pi_{33}^3 \pi_{43}^1 - \pi_{43}^3 \\ 0 \end{bmatrix} \begin{bmatrix} \beta \\ \gamma \\ \delta \\ \varepsilon \end{bmatrix} &= \alpha \begin{bmatrix} \omega_3^3 - \omega_3^2 \\ \omega_3^2 - \omega_3^1 \\ -\pi_{12}^3 \\ -\pi_{21}^3 \end{bmatrix}, \\ \begin{bmatrix} \pi_{31}^3 & \pi_{41}^3 & 0 & 0 \\ 0 & 0 & \pi_{32}^3 & \pi_{42}^3 \\ \pi_{33}^3 & \pi_{43}^3 & -\pi_{33}^1 & -\pi_{43}^1 \\ \pi_{33}^2 & \pi_{43}^2 & \pi_{33}^3 & \pi_{43}^3 \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \\ \nu \\ \tau \end{bmatrix} &= (\alpha\omega_3^1 + (\delta - \beta)\pi_{33}^1 + (\varepsilon - \gamma)\pi_{43}^1) \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix}, \end{aligned} \quad (44)$$

where $\omega_3 = \pi_{23} - \pi_{13}$ in the *original* coordinate system. The matrix from the first system is invertible, since with elementary column operations it reduces to

$$\begin{bmatrix} 0 & 0 & \pi_{33}^2 - \pi_{33}^3 \pi_{43}^2 - \pi_{43}^2 \\ 0 & 0 & \pi_{33}^1 - \pi_{33}^3 \pi_{43}^1 - \pi_{43}^1 \\ \pi_{33}^2 \pi_{43}^2 & 0 & 0 \\ \pi_{33}^3 \pi_{43}^3 & \pi_{33}^1 & \pi_{43}^1 \end{bmatrix} \quad (46)$$

which is invertible because of the general position condition. The matrix in equation (45) is also invertible, since using elementary row operations it reduces to

$$\begin{bmatrix} \pi_{32}^3 & \pi_{42}^3 & 0 & 0 \\ \pi_{33}^3 & \pi_{43}^3 & 0 & 0 \\ 0 & 0 & \pi_{31}^3 & \pi_{41}^3 \\ 0 & 0 & \pi_{33}^3 & \pi_{43}^3 \end{bmatrix}. \quad (47)$$

Thus, for every value of α , we obtain a unique solution $[\alpha, \beta, \gamma, \delta, \epsilon, \lambda, \mu, \tau]$. We now need to show that the matrix Q defined by this solution represents an actual homography. Since the determinant of Q is $\alpha(\lambda\tau - \mu\nu)$, and we can assume that $\alpha \neq 0$, we need to show that $\lambda\tau - \mu\nu \neq 0$. Indeed, from the form of (41) we note that

$$\begin{aligned} \begin{bmatrix} \lambda & \mu \\ \nu & \tau \end{bmatrix} \begin{bmatrix} \pi_{31}^3 \\ \pi_{41}^3 \end{bmatrix} &= \kappa_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \begin{bmatrix} \lambda & \mu \\ \nu & \tau \end{bmatrix} \begin{bmatrix} \pi_{32}^3 \\ \pi_{42}^3 \end{bmatrix} &= \kappa_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned} \quad (48)$$

for $\kappa_1, \kappa_2 \neq 0$ (or the retinal planes would pass through the pinholes). Finally, since in the new coordinate vectors we have $\pi_{41}^3 \neq 0, \pi_{32}^3 \neq 0, \omega_1 (= \omega_2 = \omega_3) \neq 0$, we can always normalize the Π_j matrices so that these values are equal to 1.

References

- N. Canterakis. A minimal set of constraints for the trifocal tensor. In *ECCV*, 2000.
- S. Carlsson. Duality of reconstruction and positioning in multiple views. In *Proc. Workshop on Representation of Visual Scenes*, 1995.
- O. Faugeras and B. Mourrain. On the geometry and algebra of the point and line correspondences between n images. Technical Report 2665, INRIA, 1995.
- D.A. Forsyth and J. Ponce. *Computer Vision: A Modern Approach*. Prentice-Hall, 2003.
- D.A. Forsyth and J. Ponce. *Computer Vision: A Modern Approach*. Prentice-Hall, 2012. Second edition.
- R. Hartley. Lines and points in three views and the trifocal tensor. *IJCV*, 22(2):125–140, 1997.
- R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, 2000.
- A. Heyden and K. Astrom. Algebraic properties of multilinear constraints. *Mathematical Methods in the Applied Sciences*, 20(13):1135–1162, 1997.
- H.C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, 1981.
- Q.-T. Luong and O.D. Faugeras. The fundamental matrix: theory, algorithms, and stability analysis. *IJCV*, 17(1):43–76, 1996.
- J.C. McGlone, editor. *Manual of Photogrammetry*. ASPRS, 2004. Fifth Edition.
- T. Papadopoulo and O. Faugeras. A new characterization of the trifocal tensor. In *ECCV*, 1998.
- J. Ponce and M. Hebert. Trinocular geometry revisited. In *CVPR*, 2014.
- J. Ponce, D.H. Marimont, and T.A. Cass. Analytical methods for uncalibrated stereo and motion reconstruction. In J.-O. Eklundh, editor, *ECCV*, volume 800 of *Lecture Notes in Computer Science*, pages 463–470. Springer-Verlag, 1994.
- J. Ponce, K. McHenry, T. Papadopoulo, M. Teillaud, and B. Triggs. The absolute quadratic complex and its application to camera self calibration. In *CVPR*, 2005.
- H. Pottmann and J. Wallner. *Computational Line Geometry*. Springer-Verlag, Berlin, 2001.
- L. Quan. Invariants of six points and projective reconstruction from three uncalibrated images. *PAMI*, 17(1):34–46, January 1995.
- C. Ressel. A minimal set of constraints and a minimal parameterization for the trifocal tensor. In *ISPRS-Comm. III Symposium, Vol. XXXIV, Part 3A*, 2002.
- A. Shashua. Algebraic functions for recognition. *PAMI*, 17(8):779–789, 1995.
- M.E. Spetsakis and Y. Aloimonos. Structure from motion using line correspondences. *IJCV*, 4(3):171–183, 1990.
- M.M. Thompson, R.C. Eller, W.A. Radlinski, and J.L. Speert, editors. *Manual of Photogrammetry*. ASPRS, 1966. Third Edition.
- P.H.S. Torr and A. Zisserman. Robust parameterization and computation of the trifocal tensor. *IVC*, 15, 1997.
- M. Trager, M. Hebert, and J. Ponce. The joint image handbook. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 909–917, 2015.
- B. Triggs. Matching constraints and the joint image. In *ICCV*, 1995.
- O. Veblen and J. Young. *Projective Geometry*. The Athanenum Press, 1910.
- J. Weng, T.S. Huang, and N. Ahuja. Motion and structure from line correspondences: closed-form solution, uniqueness, and optimization. *PAMI*, 14(3):318–336, 1992.